

Wavelet Encoding and Variable Resolution Progressive Transmission

Ronald P. Blanford
TRW Systems Integration Group
One Space Park, Bldg R2/2162
Redondo Beach, CA 90278

Abstract

Progressive transmission is a method of transmitting and displaying imagery in stages of successively improving quality. The subsampled lowpass image representations generated by a wavelet transformation suit this purpose well, but for best results the order of presentation is critical. Candidate data for transmission are best selected using dynamic prioritization criteria generated from image contents and viewer guidance. We show that wavelets are not only suitable but superior when used to encode data for progressive transmission at non-uniform resolutions. This application does not preclude additional compression using quantization of highpass coefficients, which to the contrary results in superior image approximations at low data rates.

1 Background

Progressive transmission is a method of encoding and transmitting imagery in such a way that gross features are able to be displayed first and subsequently refined to higher and higher resolution. Among the many possible encoding techniques are multiresolution pyramids, discrete cosine transforms, vector quantization, and wavelet transforms. Tzou [6] provides a comprehensive review of proposed techniques for progressive transmission.

The order in which the image data is selected, transmitted, and presented to the user may be dynamically prioritized as a function of both image content and immediate user interest. This typically results in a display which has a non-uniform resolution. Regions containing visually or operationally significant information may be rendered at a much higher resolution, with refinement deferred for areas of uniform intensity or lesser importance. Dreizen [3] proposed one such implementation in which the transmitter identified significant regions and communicated this information to the receiver in addition to the image data. Blanford [2] observed that for a large variety of images this overhead was unnecessary because the receiver could make a reasonable guess at the location of significant regions from image information already transmitted and displayed.

Recent results of compression using wavelet encoding have been shown to provide efficient bit rate reduction while maintaining quite acceptable image quality. The multiresolution nature of the wavelet transform, described by Mallat [5], and its computational efficiency make it a good candidate for fine-grained progressive transmission as well. Antonini *et al.* [1], for example, present

a coarse-grained example of progressive transmission using wavelets by the simple expedient of displaying each lowpass approximation as it is generated during the course of decoding.

In this paper we show that wavelets are not only suitable but superior when used to encode data for fine-grained progressive transmission at non-uniform resolutions. We first describe the approach, then discuss issues and problems in the incorporation of wavelet encoding. We present results showing a marked improvement in the approximations generated for equivalent amounts of data transmitted. Finally we show that the compressible nature of wavelets is not lost in this application; to the contrary, compression by quantization of highpass coefficients results in superior image approximations at low data rates.

2 Approach

In a prior publication [2] we presented arguments which led to the conclusion that, in the spatial domain, the low-resolution image approximation which minimizes the mean square error consists of a collection of disjoint regions each of which is painted with the average value of the pixels subsumed. For ease of computation and representation, these regions are restricted to representing nodes in a quadtree constructed by iteratively averaging groups of four pixels. An image approximation, therefore, corresponds to an arbitrary cut through the quadtree, with the minimal approximation being the global average represented by the single node at the apex. The progressive transmission of Antonini *et al.* [1], for example, can be characterized as displaying a set of horizontal cuts corresponding to uniform levels of resolution.

But the cuts need be neither horizontal nor planar. In actuality, the process of transmission may be envisioned as a walk through this quadtree. At each step in the traversal, an unvisited node is selected and expanded by transmitting the information required to generate its children. At the receiving end, the current approximation is transformed into its successor by using the new information to generate and paint the values of the child regions. Thus each approximation differs from its predecessor in that a single region has been replaced by four subregions. The traversal terminates when all regions are leaf nodes one pixel in size.

The question then arises which of the many possible traversals is optimal. For a non-interactive transmission the goal might be to minimize the mean square error, in which case a greedy algorithm can be applied at each step to select for refinement the region with maximum error. In an interactive situation, if the viewer has indicated a particular point of interest in the image then the traversal might select the non-leaf region nearest that point. If a particular feature is of interest, then the region could be selected which responds most strongly to a feature detection algorithm. Or the selection might be based on a combination of several criteria.

In this discussion we will model a non-interactive session with the goal of minimizing the error represented in the approximation. The region with the greatest error is that whose product of pixel variance and area is maximized. The receiver knows the area but not the pixel variance of the regions in its approximation. The transmitter knows both and could send a region identifier along with the information needed to refine the region, but this overhead is generally unacceptable in the low-bandwidth situations where progressive transmission is most useful. It turns out, however, that for a wide variety of images a good predictor of the pixel variance within a region is the variance between the region and those which neighbor it in the current approximation. The receiver and transmitter can independently perform this computation to select the region and only the corresponding image

information need be transmitted.

3 Image encoding

In the previous work, we introduced the additional constraints that the encoding method used to build the quadtree be lossless and introduce no storage overhead. These constraints led to the selection of the *comp/diff* encoding scheme first proposed by Knowlton [4]. In this paper we will relax first the overhead constraint and then the lossless constraint and show by comparison how the resulting approximations fare.

Knowlton’s *comp/diff* encoding applied to two pixel values produces a composite value which approximates the average and a differentiator value which approximates the difference. Each of these values requires precisely the same number of bits as the original pixels, and so requires no storage overhead. The same encoding function applied to the composite and differentiator returns the original pixels, so the procedure is lossless. All is not rosy, however, as the encoded composite value may differ significantly from the true average. The error is exacerbated if quantization is attempted in an effort at compression.

The Knowlton encoding is a non-linear function but resembles a wavelet transform with two taps: low frequency information is captured in the composite while high-frequency information resides in the differentiator. The current experiment replaces that encoding with a wavelet of eight taps. The highpass coefficients generated by the wavelet transform require roughly the same number of bits for representation as the original input, but the lowpass coefficients typically require one additional bit. Thus the encoding is not without storage overhead, which empirical evidence shows can be as high as twenty percent. Most of the additional bits are used to represent the top levels of the quadtree which are transmitted first, so the number of coefficients transmitted will be fewer than with the same amount of data using the Knowlton encoding. Our hope is that the quality of the coefficients will more than compensate for the lesser number.

Constructing the encoded quadtree presents no undue difficulties. We choose to treat the image in a toroidal manner, wrapping left to right and top to bottom, so that it will not be necessary to add extra rows and columns of padding at the lower resolutions. We use a separable wavelet function having one-dimensional 8-element low-pass and high-pass kernels.

Because of the larger basis we can no longer construct the entire quadtree, but must stop at the second (4x4) level below the apex. These sixteen lowpass values are transmitted and displayed as the initial approximation. For each region we compute a refinement priority which is the product of the region area and its external variance. The area is just the number of pixels the region represents. The external variance of region R is the mean square difference between the region value p_R and the values of regions found in parts of the displayed approximation immediately adjacent to the region. Let us compute an estimate v_R^s of the variance using those neighbors, where w_N is the length of the side shared with neighbor N and p_N is its displayed value.

$$v_R^s = \frac{1}{4w_R} \sum_{N \in \text{neighbors}(R)} w_N (p_R - p_N)^2 \quad (1)$$

After assigning priorities to all initial regions, we enter the refinement phase.

4 Image refinement

The refinement proceeds in three steps, iterated repeatedly until transmission is complete.

1. Select the maximum priority region for refinement.
2. Transmit the encoded data needed to produce the four subregions.
3. Compute priorities for the new subregions and their immediate neighbors.

The selection of the maximum priority region is simple at the beginning of the transmission, but with a brute-force approach would quickly grow intransigent as the number of regions multiplies to a significant fraction of the image size. We deal with this problem by creating and maintaining a priority heap which holds all unvisited regions. The effort to insert and update priorities in a heap with N entries is of order $O(\log N)$, which renders the problem manageable. The next region to be selected is always at the top of the heap.

With the Knowlton encoding, the region value together with three additional differentiators was all that was needed to compute the values of its four children. The broader basis of the wavelet encoding necessitates that, if region R is to be refined, then not only must its value be present but also the lowpass values of neighbors in a 5×5 area surrounding it. If a neighbor is found whose value has not yet been computed, its parent is selected for refinement regardless of priority and its 5×5 neighborhood examined for data availability. Because the initial regions all satisfy this neighborhood criterion, the procedure must eventually succeed in selecting a region.

The second step is to identify and transmit the additional information needed to produce the four children. Just as with the Knowlton encoding, we require three highpass coefficients for each lowpass one. The difference is that we now require 25 times as many, and that some may have already been transmitted for use with other regions. The bookkeeping required to determine which coefficients remain to be transmitted is painstaking but not unduly taxing. Once all required coefficients have been provided, the effort to decode the coefficients and produce the subregion values is trivial. These four values then replace the original region value in the displayed approximation.

Finally priorities must be computed for the newly created regions. Also, since these new region values contribute to the priority computations of their immediate neighbors, their priorities must be invalidated and recomputed as well. The replaced region is removed from the priority heap and all new priorities inserted or updated. This process adjusts the heap so that the new maximum priority moves to the top, in preparation for the next iteration.

5 Results of wavelet encoding

Figure 1 shows an aerial view of a portion of Moffet Field, California. The original image contains 8-bit data, 512 pixels on an edge. There are many small features as well as sharp edges between foreground and background, which make it a rather difficult image to compress effectively. Figure 2 gives a graphic comparison of the approximation error as the transmission progresses. The horizontal axis measures the number of bits transmitted as a percentage of the original image size. Knowlton encoding is shown in dark gray. Wavelet encoding is shown in black. The graph clearly

shows that wavelet encoding provides approximations with half the error over most of the transmission, a significant improvement. The light gray curve is the result of transmitting the wavelet coefficients at a uniform resolution: a breadth-first traversal of the quadtree. The knees in the curve correspond to resolution changes. The variable resolution approach is clearly superior.

It is difficult to appreciate the actual impact of progressive transmission in a static presentation such as this. Figures 3 and 4 show snapshots of the display as it would appear when 2% and 5% of the data has been transmitted, using the wavelet encoding. Figures 5 and 6 show the corresponding displays using the Knowlton encoding. The Knowlton encoding provides higher contrasts and sharper edges where they are found, but the wavelet encoding provides a more balanced development.

Figures 7 and 8 show results of transmission error for an aerial view of the Los Angeles airport. The image itself is hazy with few contrasts so it should exhibit lower overall levels of error, as the graph bears out. The black curve for variable resolution wavelet encoding again shows half the error of the dark gray Knowlton encoding. The light gray curve for breadth-first transmission is close to that for variable resolution, as one would expect when few features stand out.

Figures 9 and 10 show results for yet another aerial view of an airport, this one a Spot satellite image of Beirut, Lebanon. Though the overall intensity and feature distribution are different, this image exhibits approximation error similar to the Los Angeles image.

6 Results of coefficient quantization

One reason wavelets provide a good basis for image compression is that the result degrades gracefully under quantization, even at extreme levels. In order to verify that this characteristic had not been lost when used with progressive transmission, a simple quantization scheme was applied to the wavelet coefficients and the impact on the approximation error observed. Briefly, this scheme divides all highpass coefficients in a given level of the quadtree by the same value. This value is greatest at the lowest level, the leaf nodes, and is reduced by a factor of two at each level above.

Figure 11 shows the results when applied to the Moffet Field image. The thick black curve on the right represents the uncompressed wavelet transmission, just as in Figure 2, though shown at an expanded scale. The second curve from the right represents a quantization factor of 4 at the base level of the quadtree, of 2 at the level above the base, and the remainder of the coefficients left unquantized. The base-level quantization factor is doubled for each successive curve, so that the leftmost curve begins with a quantization factor of 64 at the base level of the quadtree, reduced successively to a divisor of 2 at level four, with only the original lowpass coefficients remaining unscathed.

While the higher levels of quantization introduce significant error in the later stages of the transmission, the initial portions critical for early identification of features show only improvement. More significantly, the smooth degradation indicates that efforts toward designing more sophisticated quantization schemes would not go unrewarded.

7 Conclusions

In this paper we have shown that wavelets are not only suitable but superior when used to encode data for fine-grained progressive transmission at non-uniform resolutions. The results show a marked improvement in the approximations generated for equivalent amounts of data transmitted when wavelet encoding is used in place of Knowlton encoding. Finally we have shown that the compressible nature of wavelets is not lost in this application; to the contrary, compression by quantization of highpass coefficients results in superior image approximations at low data rates.

References

- [1] M. Antonini, M. Barlaud, P. Mathiew, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1:205–220, 1992.
- [2] Ronald P. Blanford. Progressive refinement using local variance estimators. *IEEE Transactions on Communications*, to appear.
- [3] Howard M. Dreizen. Content-driven progressive transmission of grey-scale images. *IEEE Transactions on Communications*, COM-35(3):289–296, March 1987.
- [4] Ken Knowlton. Progressive transmission of grey-scale and binary pictures by simple, efficient, and loss-less encoding schemes. *Proceedings of the IEEE*, 68(7):885–896, July 1980.
- [5] S. Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.
- [6] Kou-Hu Tzou. Progressive image transmission: A review and comparison of techniques. *Optical Engineering*, 26(7):581–589, July 1987.

Moffet Field data:



Figure 1: Moffet Field original image

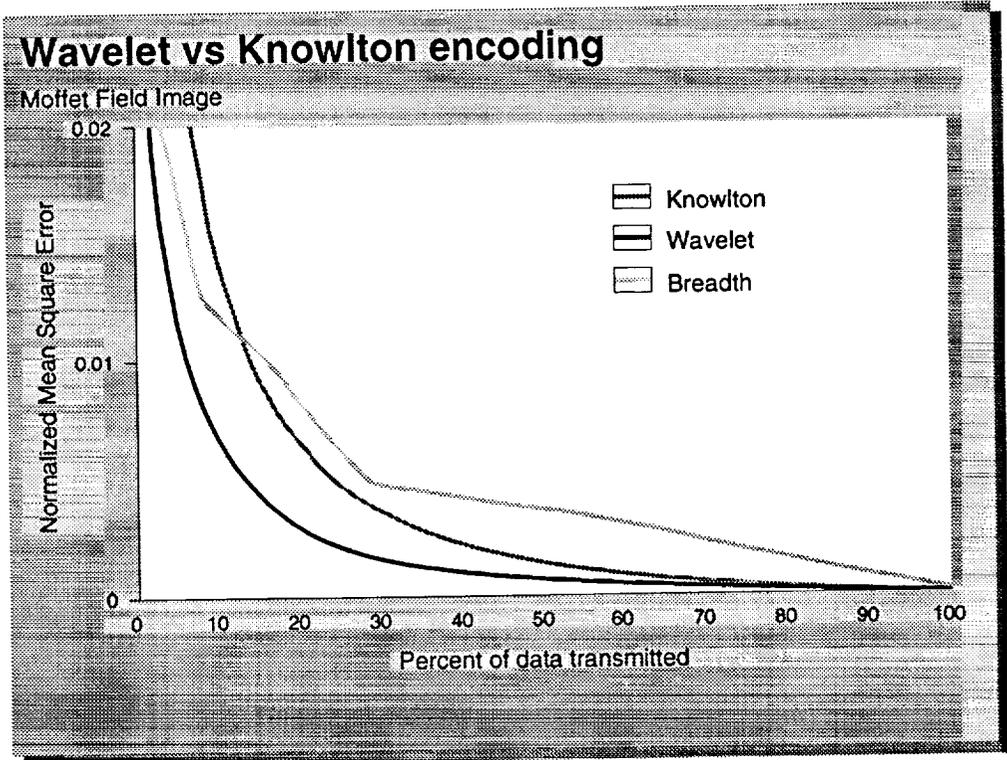


Figure 2: Comparison of encoding error

Moffet Field snapshots:



Figure 3: Wavelet encoding, 2% of data

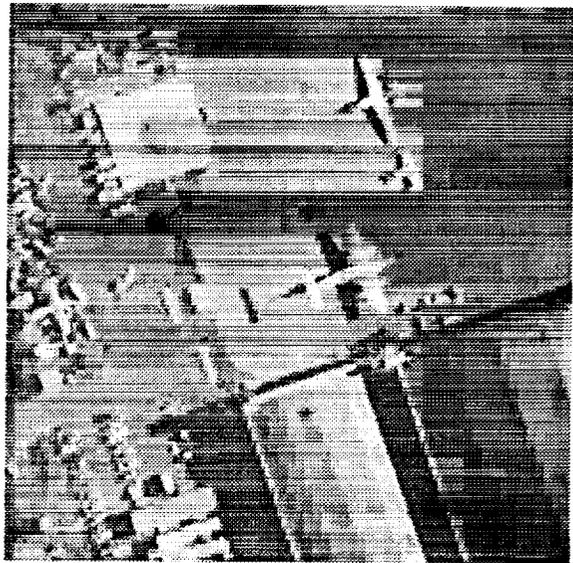


Figure 4: Wavelet encoding, 5% of data



Figure 5: Knowlton encoding, 2% of data

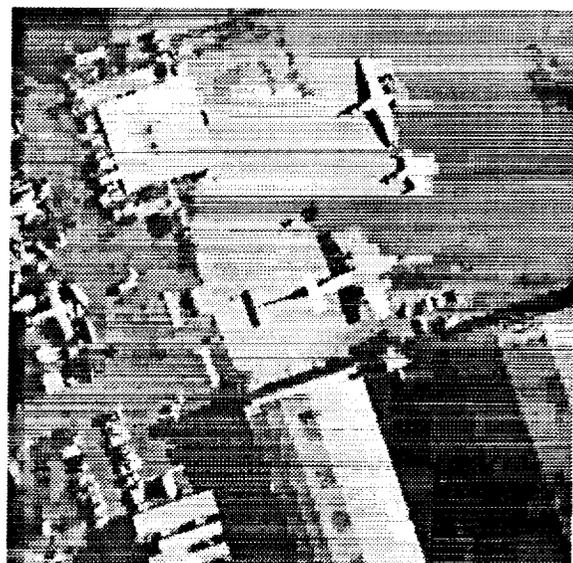


Figure 6: Knowlton encoding, 5% of data

Los Angeles data:

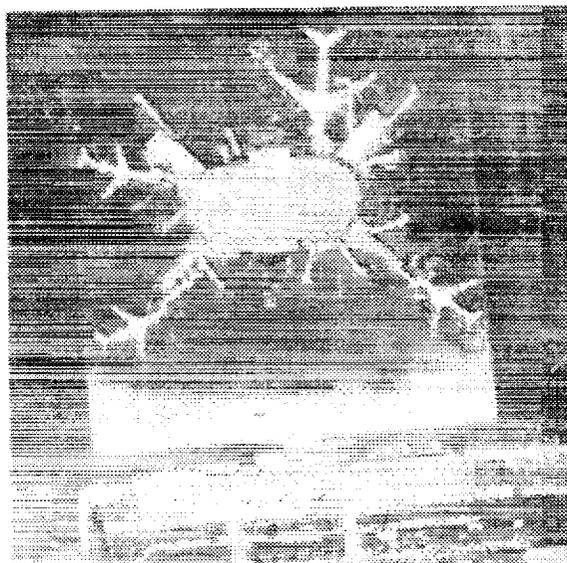


Figure 7: Los Angeles Airport original image

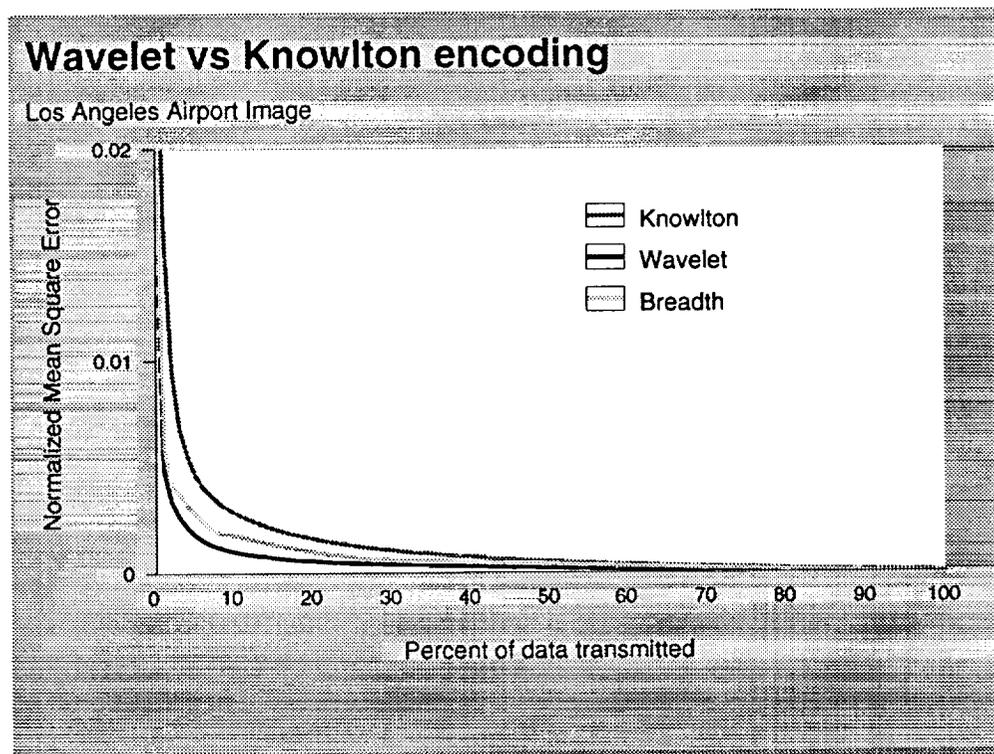


Figure 8: Comparison of encoding error

Beirut Airport data:

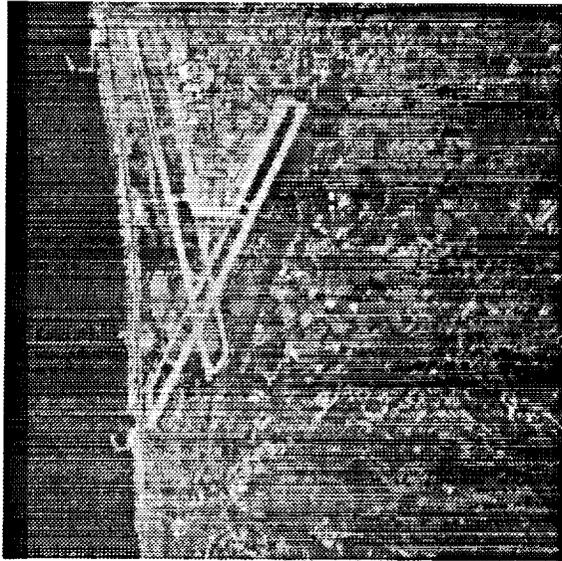


Figure 9: Beirut Airport original image

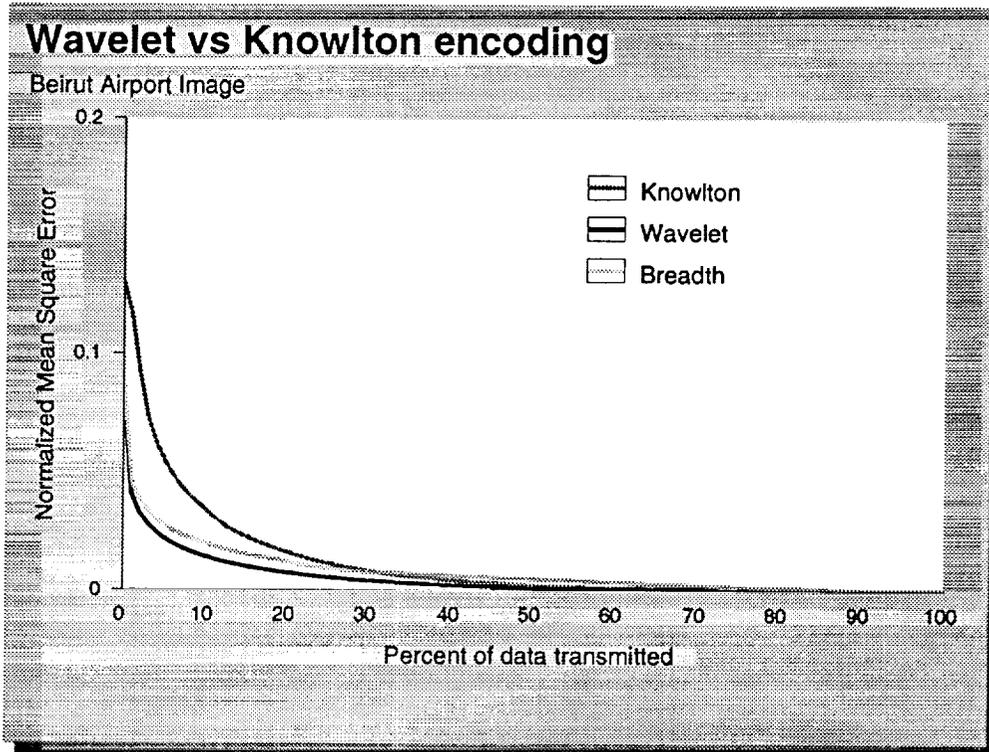


Figure 10: Comparison of encoding error

Quantization data:

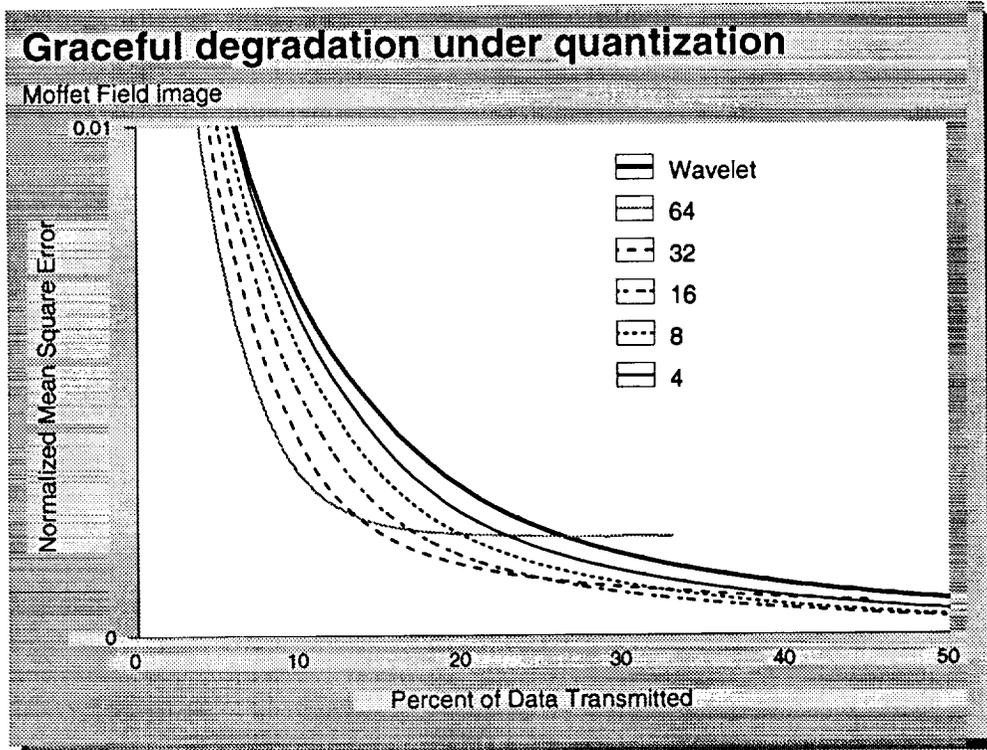


Figure 11: Comparison of quantization error

